Value at Risk (VaR) Backtesting Analysis

1. Introduction

Value at Risk (VaR) is integral to market risk management in financial institutions. It is particularly relevant given the increasing market volatility and regulatory requirements in the past few years. The following analysis assesses the performance of three distinct VaR estimation methods (Historical Simulation, Normal GARCH, and Student-t GARCH). These methods will be applied to three major banking institutions, namely: JPMorgan Chase (JPM), Barclays (BCS), and Deutsche Bank (DB). We have chosen these firms as they represent different geographical markets (US, UK, and European) and faced varying degrees of market stress during the studied period. This allow us to study effectively our different methodologies under diverse market conditions.

After implementation, we backtested our three VaRs to assess their reliability in terms of market risk prediction. In the second part, we extracted time-varying student-t GARCH parameters to understand how risk dynamics changes over time. The analysis uses daily return data from WRDS that spans January 2019 to December 2023. This period shows both normal and stressed market conditions. Please note that the first 100 days of 2019 are not plotted as they represent our first values for the rolling window.

2. VaR Backtesting Methodology

Our analysis implements three distinct Value-at-Risk methodologies. Each of these are estimated using a rolling window in order to adapt to the dynamism of market risk. The models are calibrated at the 1% confidence level with a portfolio value scaled to 1, allowing for direct comparison of results.

2.1 Model Specifications

Historical Simulation (HS-VaR) The Historical Simulation represents a non-parametric approach using a 100-day rolling window. For each time point t, the VaR is calculated as the empirical 1st percentile of the previous 100 daily returns. Because of its reliance on the rolling window, this simulation shows a step-like pattern. This consequently makes it slower to adapt extreme events. It makes no distributional assumptions about returns, instead relying purely on historical data to estimate potential losses.

GARCH Models We use a GARCH(1,1) specification with zero mean to reduce parameter estimation noise. The models are re-estimated daily using the rolling window to capture time-varying volatility. We test two distributional assumptions in our analysis. The Normal GARCH assumes that unpredictable component of the return is normally distributed. The parameters are estimated via maximum likelihood and VaR is calculated using the normal quantile at 1%. The Student-t GARCH extends this framework by incorporating a heavy-tailed distribution, with an additional degrees of freedom parameter estimated daily and VaR calculated using the Student-t quantile at 1%.

2.2 Backtesting Framework

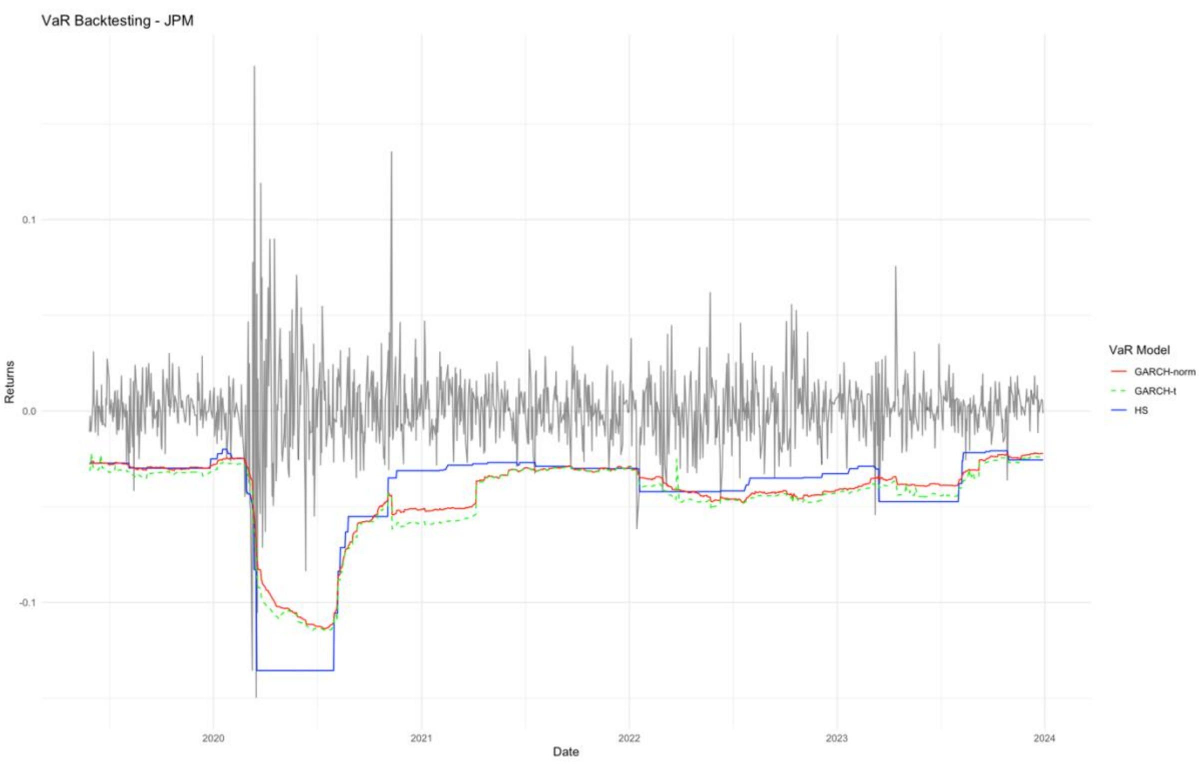
The backtesting process checks violations, precisely, where actual returns fall below VaR estimates. The violation rate is computed as follow: number of violations divided by the total number of observations. It is then compared to the expected rate of 0.01 at our chosen 1% confidence level. Additionally, we performed a Kupiec test per VaR. It's null hypothesis says that the observed violation rate is equal to the expected rate. Rejection of this hypothesis at the 5% significance level indicates model misspecification.

2.3 Graphical Analysis

The backtesting plots provide a visual representation of our results, displaying daily returns (grey line) against which we can compare our VaR estimates. The Historical Simulation estimates are shown in blue, Normal GARCH in red, and Student-t GARCH in green. A violation occurs when the return line crosses below any of the VaR estimate lines. This method allows us to observe not only the frequency of violations but also their timing and their severity.

2.4 JPMorgan Chase (JPM)

<u>A. Context and Market Analysis:</u>



As seen in the above plot in grey, JP Morgan experienced its highest volatility spike in March 2020 when the COVID-19 pandemic caused a global market crash. As a result, businesses shut down around the world and the bank consequently faced risks of possible loan defaults. Also, the Federal Reserve's emergency rate cuts during this period directly impacted the banking sector profitability. However, through strong risk management the company showed resilience and even benefited from increased trading revenue during the market recovery in late 2020. This later lead to a more stable period in 2021.

B. Results and Backtesting Interpretation:

Violation Rates:

Historic Simulation VaR	Standard GARCH	Student-t GARCH
0.01987281	0.02146264	0.01589825

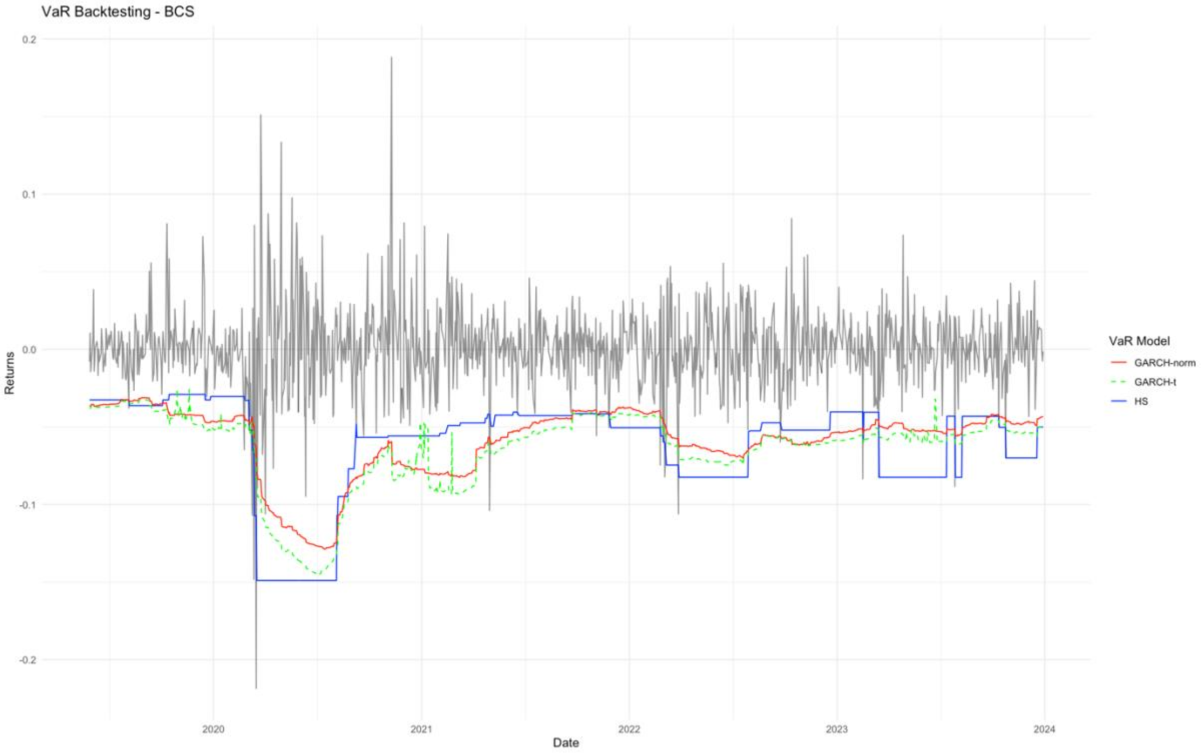
Kupiec Test p-values:

Historic Simulation	Standard GARCH	Student-t GARCH
0.0005930939	0.0001029160	0.0242018963

The backtesting results for JPM indicate a systematic underestimation of risk across all three models. The Student-t GARCH performs best with a violation rate of 1.59% (20 violations), though still above the expected 1% threshold. Both Normal GARCH and Historical Simulation show higher violation rates at 2.15% and 1.99% respectively. The Kupiec test rejects all models at the 5% significance level (p-values: 0.024 for Student-t GARCH, 0.0001 for Normal GARCH, and 0.0006 for HS). The rejection of the null hypothesis implies that the actual violation rate is statistically different from the expected rate. The high violation rates correspond with the periods of heightened market stress. As we can see in the above plot, this is particularly shown during the March 2020 COVID-19 market crash, where all models struggled to capture the extreme market movements. These extreme movements violate our models assumptions which results in such backtesting statistics.

2.5 Barclays (BCS)

A. Context and Market Analysis:



As shown in the plot, Barclay's volatility, as JP's, spikes in March 2020 when COVID-19 impacts were amplified by UK-specific concerns. During this period, the Brexit had already made UK banks particularly vulnerable. This resulted in London's future as a financial centre being questioned and added market stress. In 2021, the bank faced legal issues due to a trading scandal and was fined $361 million which can be perceived in the graph. Over most of the studied period, UK regulatory changes and the medium-term impact of Brexit created ongoing volatility, unique to British financial institutions.

B. Results and Backtesting Interpretation

Violation Rates:

Historic Simulation VaR	Standard GARCH	Student-t GARCH
0.01748808	0.01828299	0.01510334

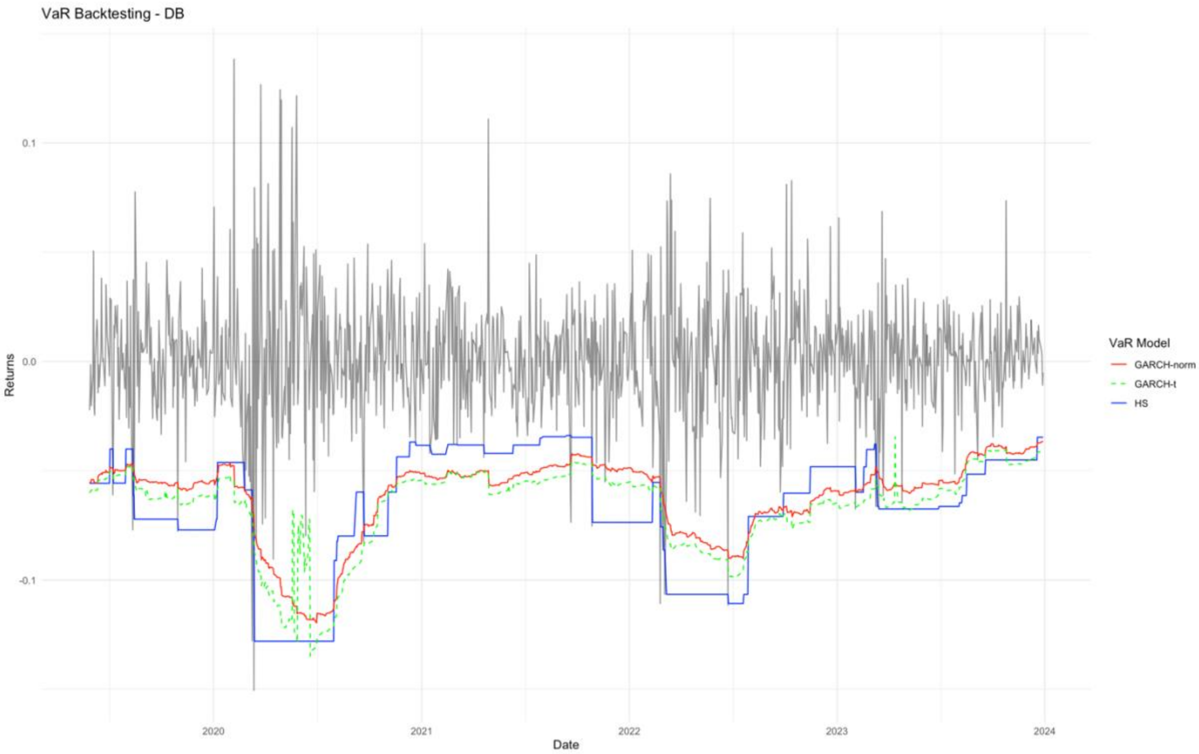
Kupiec Test p-values:

Historic Simulation	Standard GARCH	Student-t GARCH
0.006195413	0.002946963	0.044852157

Barclays shows similar patterns but with slightly better overall performance. The Student-t GARCH again emerges as the most reliable model with a violation rate of 1.51% (19 violations), followed by Historical Simulation at 1.75% and Normal GARCH at 1.83%. However, the Kupiec test still rejects all models at the 5% level (p-values: 0.045 for Student-t GARCH, 0.003 for Normal GARCH, and 0.006 for HS). The rejection is less severe than for JPM, especially for the Student-t GARCH model. This suggests somewhat a better risk capture even when considering the tough UK market conditions. The violations occurs around periods of UK-specific market stress, which includes the challenges described above.

2.6 Deutsche Bank (DB)

A. Context and Market Analysis:



Deutsche Bank had its most extreme volatility during March-April 2020, with the COVID-19 impact particularly severe. At the same time, the bank was having a major restructuring process. DB faced additional pressure in 2022 because of the European energy crisis. This followed from the Russia-Ukraine conflict and specifically impacted the German banking sector. Throughout this period, Deutsche Bank's volatility patterns reflected both global market conditions and specific European economic challenges, including significant ECB interest rate policy changes.

B. Results and Backtesting Interpretation

Violation Rates:

Historic Simulation VaR	Standard GARCH	Student-t GARCH
0.01669316	0.01828299	0.01430843

Kupiec Test p-values:

Historic Simulation	Standard GARCH	Student-t GARCH
0.012505028	0.002946963	0.079487881

Deutsche Bank demonstrates the best results in the implementation among the three institutions. The Student-t GARCH achieves the lowest violation rate at 1.43% and, notably, is the only model across all banks and methods to pass the Kupiec test at the 5% significance level (p-value: 0.079). The Normal GARCH and Historical Simulation show violation rates of 1.83% and 1.67% respectively, with both models rejected by the Kupiec test (p-values: 0.003 and 0.013). These results suggest that the Student-t GARCH capturing heavy tails is very effective for DB's return distribution, possibly due to the bank's exposure to both European systematic risk and internal factors during its restructuring period.

2.7. Conclusion on Observations

The analysis of log returns across all three banks reveals important similarities in market behaviour. The most pronounced pattern is the strong volatility clustering observed during March 2020. This event triggered extreme market stress across global financial markets, reflected in sharp increases in returns volatility for JPMorgan Chase (JPM), Barclays (BCS), and Deutsche Bank (DB). While each bank faced unique regional challenges—JPM with Federal Reserve policy shifts, BCS with Brexit-related uncertainties, and DB with the European energy crisis—their fundamental risk patterns showed high correlation during major global market events. Post-2021, gradual stabilisation in returns volatility was observed, driven by a combination of government interventions and the banks' adaptive measures to the post-pandemic economic landscape.

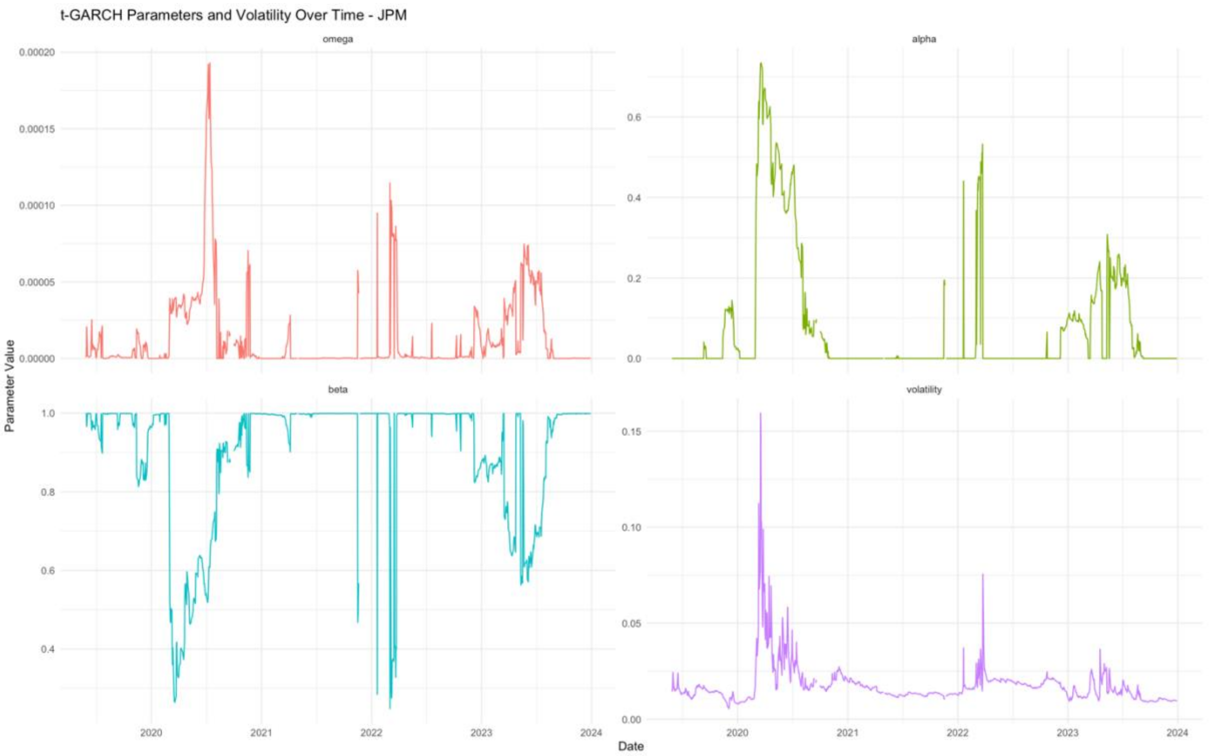
The comparative results highlight the superior performance of the Student-t GARCH model across all three institutions, though with varying degrees of success. The consistently higher violation rates of the Normal GARCH model underscore the importance of accounting for heavy tails in financial return distributions. The Historical Simulation's performance generally falls between the two GARCH models, suggesting that while it avoids distributional assumptions, its rigid dependence on the historical window limits its effectiveness during rapidly changing market conditions.

3. JP’s Student-t GARCH parameters extraction

JPMorgan Chase (JPM) was chosen for parameter analysis due to its positioning as the largest U.S. bank but also because of its important exposure to global market volatility. Although our model didn't perform as well as for the other institutions, JP gives us an interesting case study for studying risk dynamics under diverse market condition. The company's operations span both retail and investment banking. This nature makes it highly sensitive to shocks in the market. The sample period (2019-2023) is ideal in terms of context to evaluate the model's ability to capture extreme market dynamics. All of these factors consequently raises the question on whether the Student-t GARCH effectively captures risk patterns for an institution as globally interconnected as JPMorgan Chase or not.

3.1 Analysis of GARCH Parameters for JPM:

Our analysis uses a 100-day rolling window to re-estimate Student-t GARCH parameters and volatility for JPMorgan Chase and is updated on a daily basis. The results illustrate the dynamic behaviour of four key components that we plotted. Omega (ω) represents baseline variance. Alpha (a) measures sensitivity to new shocks. Beta (ß) captures the persistence of volatility. Realised volatility measures overall market risk. The analysis of these components allows us to understand how risk can evolve over time. During normal conditions, the parameters show a certain stability, with high beta reflecting persistent volatility and low alpha indicating limited sensitivity to shocks. On the other hand, during periods of market stress, (e.g. covid), we observe distinct shifts. Omega spikes, indicating heightened baseline risk, while alpha increases sharply, reflecting greater sensitivity to market shocks. Beta, which is typically near 1, temporarily declines during these periods, showing that long-term volatility patterns are disrupted. The clustering and spikes of the market event described for JP, is further highlighted by the realised volatility plot. This reinforces the model's ability to adapt in a dynamic manner to changing market conditions.



3.2 Conclusion on observations

The Student-t GARCH model is, as proved above with its adaptive degrees of freedom, very effective in capturing the kurtosis of financial returns. Its ability to adapt to shifts in risk conditions makes it very valuable for understanding extreme market dynamics. The parameters fluctuations, like the numerous spikes in omega and alpha, gives valuable informations for risk management. These could be, for example, useful in asset reallocation and hedging strategies. The persistence of volatility, reflected in beta, shows the need for a certain sustained caution even as instability arrives in the market. However, this analysis has some limitations. Firstly, the high variability in the different parameter estimates during stressed period, clearly suggests some instability when estimating. Secondly, While the 100-day rolling window improves the responsiveness to changing market conditions, it may introduce noise particularly in omega and alpha. Furthermore, the simple structure of the GARCH(1,1), while computationally efficient, does not completely capture the complexity of market dynamics.

Despite these challenges, the Student-t GARCH framework offers very important informations into JPMorgan's risk profile. It allows us to understand the importance of adopting more flexible and dynamic models for managing risk in volatile markets.

4. Implications and Recommendation

The findings give us considerations for risk management and model selection across varying market conditions:

Historical Simulation seems to be the most effective model in stable markets because of its simplicity. However, it has a certain rigidity during volatile periods. This consequently put limitation on its reliability as a standalone model.

Normal GARCH, while computationally efficient and responsive to changing volatility, it tends to underestimate tail risks. Our implementation also suggests adding safety margins or other models for regulatory compliance.

Student-t GARCH appears as the most reliable model during complicated times. It takes into account tail risks and extreme events. Its strong performance across all three institutions, particularly for Deutsche Bank, proves its value in real-world risk management and regulatory applications.

Our findings put emphasis on the importance of having a flexible and consistent risk framework which has adaptability during crises. In our opinion, the best alternative would be to build hybrid models that combine both parametric and non-parametric approaches. We could also extend our analyses to multi-asset portfolios, add macroeconomic indicators, or for example, sentiment analysis from social media in our model in order to be more accurate. In summary, for risk management in volatile markets, the Student-t GARCH model gives us the most informations. Its ability to reflect evolving market conditions and account for heavy tails makes it very valuable. Through this, financial institutions can optimise capital allocation, and have a better understanding of extreme market stress.